DESIGNING PRIMERS TO TEST FOR DENITRIFICATION BIOMARKERS IN SEPTIC SYSTEM LEACH FIELDS

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the

Master of Engineering

by

Selene Ka-Wan Leung

August 2016

ABSTRACT

Septic sewer systems are an increasingly popular method for wastewater treatment in domestic households across the United States. Investigating their greenhouse gas emissions and the microbial communities controlling the cycling of the two most common greenhouse gases, nitrous oxide (N_2O) and methane (CH_4), provides important information on their potential atmospheric impacts and helps streamline the use of molecular biology techniques to correlate microbial activity to net greenhouse gas fluxes. Over the course of a summer, gas flux measurements and soil samples from the leachfields and control plots of nine different septic systems in central New York were analyzed. While preliminary gas flux data and quantitative polymerase chain reaction (qPCR) data suggests that soil depth is a greater predictor of denitrifier activity than the presence of a septic leachfield, more investigation on the potential quantity, roles, and types of denitrifiers must be conducted. Thus degenerate forward and reverse primers targeting the biomarker genes of *nosZ* and *norB* as well as their appropriate protocols were designed, validated, and optimized using an array of biomolecular techniques such as PCR, gel electrophoresis, qPCR, and high-throughput sequencing. Long amplicon forward and reverse primers and their design protocols were also created, validated, and optimized in order to amplify the limited quantity of genetic material extracted from the soil samples without depleting original sources. Standard curves of these degenerate primers are being created and optimized for further in-depth analysis of the exact abundance of selected denitrifiers in the soil samples. Both the end-point and long amplicon primers will be additionally used to help determine the microbial community composition at the septic system leachfields involved in greenhouse gas cycling. The results from this study will help inform future mitigation methods for potential greenhouse gas emissions from septic system leachfields.

BIOGRAPHICAL SKETCH

Originally from Fremont, California, Selene Ka-Wan Leung earned her Bachelor of Science in Biological Engineering with a minor in Biomedical Engineering from Cornell University in May 2015. She continued her studies at Cornell University, earning a Master of Engineering degree in Biological and Environmental Engineering in May 2016. In her spare time, she enjoys reading great literature, writing, drawing, and hiking.

ACKNOWLEDGEMENTS

I wish to first thank Professor Todd Walter, who has advised and supported me throughout five years in the Soil and Water Lab of the Biological and Environmental Engineering Department. He has enabled me to differentiate my experiences and skills, as well as imbued me with the sense of positivity and perseverance that all researchers should have when overcoming difficulties. I will forever be thankful that he believed in my potential, because he changed my life the day he walked in late to the Professor-Student dinner at Appel Commons my freshmen year (wearing his infamous tie-day shirt) and enthusiastically agreed to allow me to work for him when I mustered the courage to ask.

I would also like to thank Cristina Fernandez-Baca, who guided and mentored me throughout the development and completion of this project. Her intelligence, kindness, and patience allowed me to comfortably maneuver my way around a new laboratory setting and learn new techniques. Additionally, I would like to thank Brian Gramlich Rahm, Professor Ruth Richardson, and Allison Truhlar for allowing me to join their project and their willingness to help even when solving the simple problems.

Lastly, I would like to thank Professor Jean Hunter, Brenda Lee Marchewka, and the Biological and Environmental Engineering staff for helping me throughout the completion of my Bachelor of Science and Master of Engineering programs at Cornell University.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

SEPTIC SYSTEMS OVERVIEW

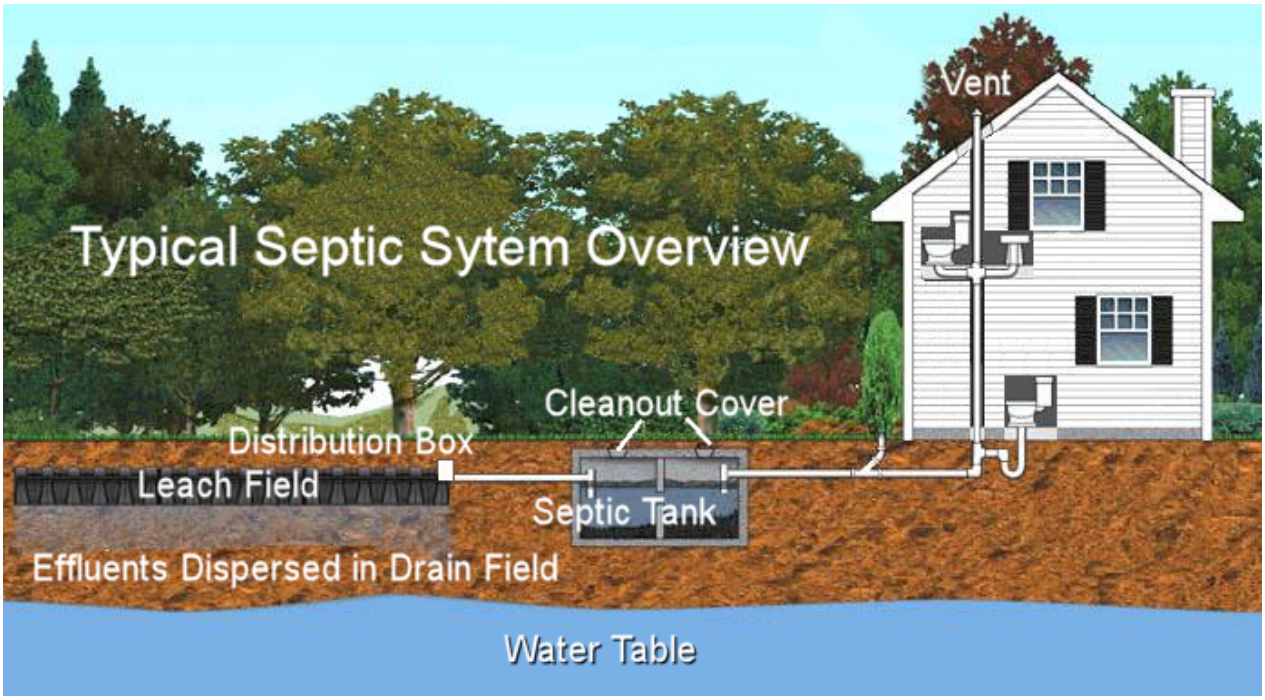

Figure 1: Domestic Septic System Overview [1].

Domestic septic systems account for 26 million individual wastewater systems in the United States and are steadily increasing in usage as homeowners recognize their sustainability, effectiveness, and easy implementation [2]. Nearly half of all houses have septic systems in the South, due to the prevalence of rural communities with limited access to a central plumbing network [3]. In addition, their long life expectancy of 25 to 30 years makes them an especially popular option in sparsely populated suburban areas where running sewer lines would be cost-prohibitive [3].

As shown in Figure 1, wastewater from bathrooms, kitchens, and laundry rooms flows into an underground septic tank, where solids and particulates are left to settle to form sludge. Vents in the septic tank allow for accumulated gases to disperse, with access ports along the

pipes. Greases and oils rise to the top of the residual liquid as scum. Effluents from the septic tank then flows into a nearby soil dispersal system, known as a leachfield [4]. Effluent from the leachfield's perforated distribution pipes gradually filters through the soil, which naturally removes harmful coliform bacteria, viruses, and nutrients, continuing to percolate through the soil until it reaches groundwater. These leachfields are often located in homeowners' backyards, or adjacent to their property, and can span from 10 feet to 30 feet depending on the size of the residence. From this leachfield, atmospheric gases are emitted as well as groundwater gases.

There is a dearth of scientific literature on the potential greenhouse gases (GHG) being emitted from these septic leachfields, as well as the microbial communities which may be controlling this GHG cycling. As of 2015, there has only been one study quantifying the GHG emissions originating from septic leachfields soils, with little knowledge on the microbial communities responsible for controlling GHG fluxes [5]. It is vital to identify the extent of the roles of these microbial communities in producing GHGs, as it allows for the future design and implementation of septic systems to be adjusted such that any GHG footprint can be mitigated, if proven to be a widespread concern. This is especially important in cases of failing or abandoned septic systems, such as when leachfields soils are oversaturated with untreated wastewater replete with excessive organic content.

The design of standard septic systems is such that any residual leachate exiting through distribution pipes is exposed to varying soil conditions of high and low water content, and acidic and basic environments. The variance in soil conditions allows for the flourishing of microbial communities comprised of methanotrophs, which are methane (CH_4) consumers, as well as aerobic and anaerobic denitrifiers, known as nitrous oxide (N_2O) producers. These microbes aid in the digestion of excessive nutrients found in the wastewater. In addition, standard septic

systems tanks support anaerobic, nutrient-rich environments which generate ideal living conditions for anaerobic methanogens, known as CH_4 producers, and denitrifiers, previously stated as N_2O producers. As a result of their prominence in the soils surrounding standard septic system leachfields, the microbial communities producing the greenhouse gases CH_4 and N_2O are of interest.

THE ROLES OF GHG PRODUCING MICROBES

Methanogens create CH_4 through methanogenesis while methanotrophs consume CH_4. Methyl coenzyme-M reductase (MCR) is an enzyme found in methanogens responsible for the final step of methanogenesis. McrA is the gene which encodes for the alpha subunit of MCR, which is correlated with CH_4 production. Thus MCR is an appropriate target in PCR based methods for detecting methanogens and a suitable biomarker for environmental applications [6,7]. Methanotrophs, or methane oxidizing bacteria, respire methane and therefore reduce the net flux of GHGs to the atmosphere. The enzyme particulate membrane-bound methane monooxygenase (pmoA) is found in aerobic, nitrite- and sulfate-reducing environments characteristic of septic systems. Because pmoA is responsible for the first step of methane oxidation, its presence is an indication of methanotroph activity [8,9,10].

The analysis of potential GHG emitters also extends to microbes responsible for denitrification, a critical step in the nitrogen cycle responsible for the reduction of nitrate (NO_3^-) to nitrogen gas (N_2). The enzyme responsible for this critical reduction step of N_2O to N_2 is nitrous oxide reductase (NOS), with the gene encoding for this step being nosZ. Because there is minimal benefit for denitrifiers such as *Paracoccus denitrificans* to completely reduce N_2O to N_2, a limited amount of bacterial species carries nosZ [11,12]. Thus, only a few studies have investigated the correlation between presence of nosZ in these bacterial species and N_2O fluxes

[13]. As shown in Figure 2, the enzyme responsible for the reduction of nitric oxide (NO) to N_2O is nitric oxide reductase (NOR), encoded by the norB and norC genes [14,15]. Thus, to analyze the full cycling of N_2O, the consumption and production of denitrification should be analyzed using their respective biomarker genes, nosZ and norB.

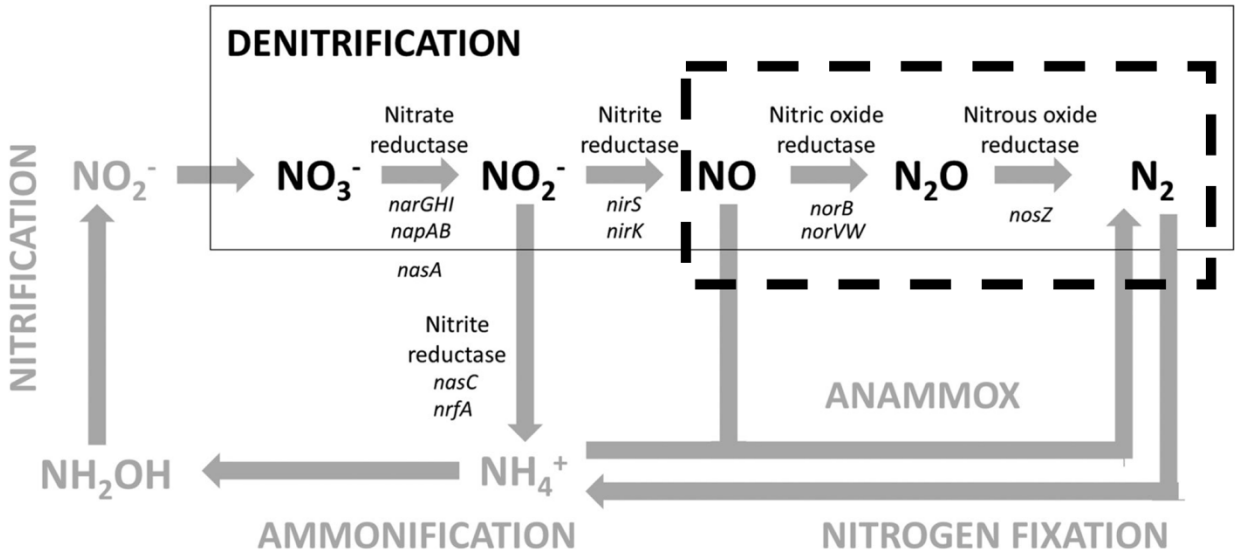

Figure 2: Denitrification Cycle and the enzymes and gases being investigated, squared in dashes [16].

EXPERIMENTAL SETUP

A preliminary analysis of the gas flux measurements of N_2O and CH_4 was conducted at nine different domestic septic systems in central New York during the summer of 2014. A set of three gas chamber collars was installed into areas identified as the site of leachfields and an additional set was installed in areas identified as the control plots, which were predominantly grassy lawns adjacent to the leachfield sites. Measurements were made at three different times during the duration of the summer at the leachfields and control lawns. Gas samples were extracted from chambers placed over the gas chamber collars via syringes and injected into evacuated gas vials. Triplicate measurements of soil moisture, temperature, and the pH were also

recorded at these sites while measurements were being taken. Gas fluxes were calculated using the gas chromatography machine, where the change in concentration of each chamber's gases were calculated over a time period of 30 minutes [17].

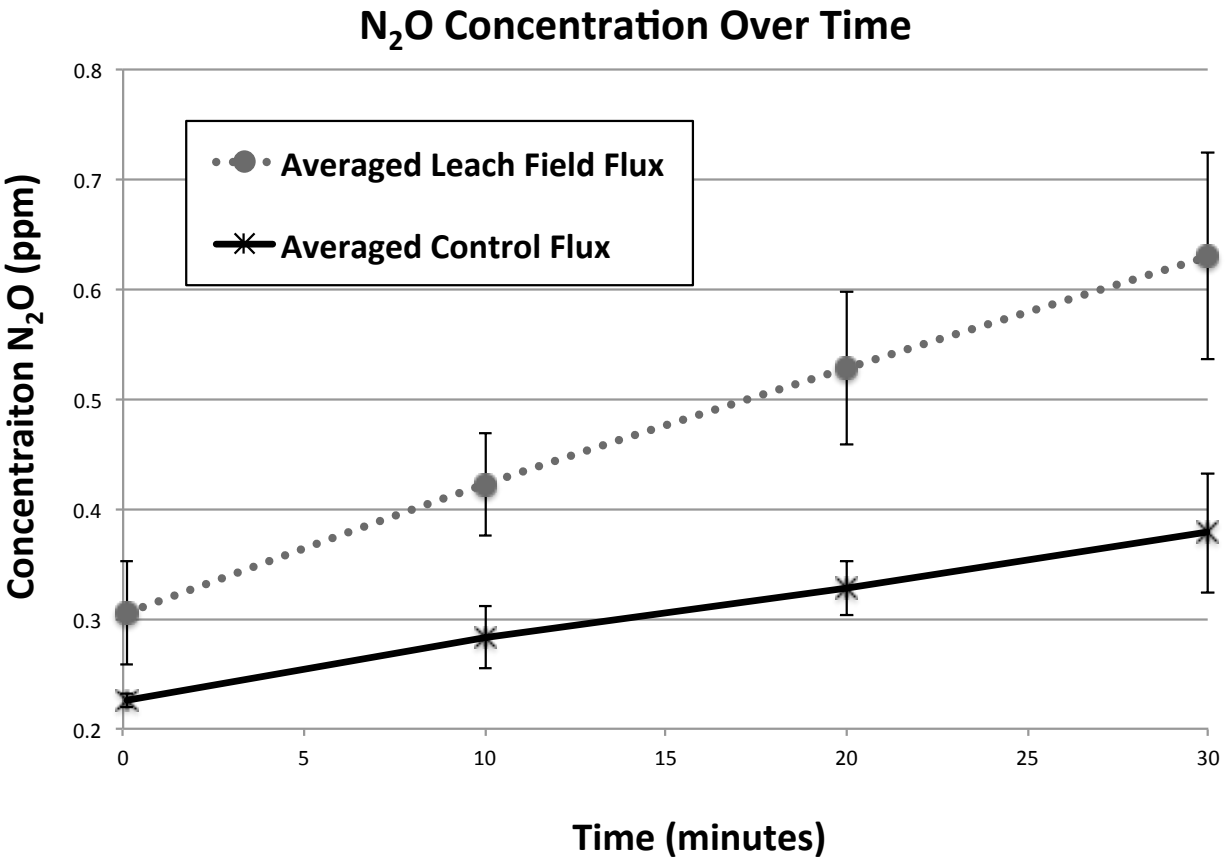

Figure 3: Change in N_2O Concentration over time of leachfield fluxes versus control field sites. Data points are the average of triplicate measurements [17].

As shown in Figure 3, preliminary results from gas analysis suggest that N_2O fluxes were greater over leachfields than control lawns. In addition, there were no meaningful differences between CH_4 fluxes over leachfields compared to the control lawns. At the same time, there was a greater total flux of methane from the septic system vents than from the leachfields, while the emissions of N_2O are similar from the vent systems and leachfields [17].

Soil samples from leachfields and control plots were also taken during days in which flux measurements were taken. Samples were collected at two different depths representing upper (0-4 inch) and lower (4-8 inch) soil layers. To test for the presence and absence of methanogens, methanotrophs, and denitrifiers, using the biomarker genes *mcrA*, *pmoA*, and *nosZ*, respectively, several soil samples were analyzed using endpoint PCR. While both leachfields and control sites have been found to contain functional genes, it remains to be detected whether the nosZ and norB genes are more abundant at which soil layer depths at the different sites [17].

EXPERIMENTAL AIMS

The implementation and analysis of this experiment is a collaborative effort between students under Dr. Ruth E. Richardson, faculty member of Cornell University's Civil and Environmental Engineering Department and students under Dr. Todd Walter, faculty member of Cornell University's Biological and Environmental Engineering Department. As such, the larger goals outlined below are the experimental aims of the overall experiment. The first goal of this experiment is to quantify the abundance of denitrifiers, methanotrophs, and methanogens from soil samples obtained from septic system leachfields. Using qPCR, the following three hypotheses will be tested: 1) There is a greater abundance of denitrifiers in the upper soil layers of septic systems compared to lower layers, since ammonia (NH_3) is converted to nitrate (NO_3-) nearer to the soil surface. 2) Methanogens are more abundant at depths closer to leachfield distribution pipes, where environmental conditions are usually nutrient rich and anaerobic. 3) Methanotrophs are more abundant in the upper soil layers, where environmental conditions are usually aerobic [17].

The second goal of this overall project would be to use high-throughput sequencing to determine the microbial community composition of septic leachfields. The biomarkers to be investigated are denitrifiers surveyed by nosZ and norB, methanogens surveyed by mcrA, and methanotrophs surveyed by pmoA. Preliminary analyses on the role of methanogens and methanotrophs have been conducted but have not yet been published [17]. The focus of this project will be to help quantify the abundance and composition of denitrifiers in order to test the aforementioned first hypothesis of the first goal and lay out the groundwork for accomplishing the second goal. These goals will be achieved through the design and testing of qPCR and PCR primers used to analyze the microbial cycling of N_2O. These primers will be targeted towards the functional biomarker genes of norB and nosZ in order to encapsulate the complete production and consumption of denitrification.

METHODS

VALIDATING DEGENERATE PRIMERS

Degenerate primer sequences targeting the denitrification pathway were obtained from previous studies which determined that the primers nosZ1 and cnorB2/B6 as shown in Table 1 were promising candidates for targeting the nosZ and norB gene, respectively [18,11,19]. The qPCR primers were ordered from Integrated DNA Technologies (IDT) and validated via end-point polymerase chain reaction (PCR) and agarose gel electrophoresis for the effectiveness of their specific binding and amplification of their PCR products [20, 21]. Gel electrophoresis will additionally provide a qualitative assessment of the amount of PCR product generated.

Primer	**5' to 3' Sequence**	**Amplicon Size (base pairs)**
nosZ1F	WCSYTGTTCMTCGACAGCCAG	248
nosZ1R	ATGTCGATCARCTGVKCRTTYTC	248
cnorB2F	GACAAGNNNTACTGGTGGT	390
cnorB6R	GAANCCCCANACNCCNGC	390

Table 1. Degenerate denitrification pathway primers for end-point PCR.

END-POINT PCR

1) Spin down primers to ensure that all of their contents are on the bottom. Resuspend the nosZ and norB primers to a 100 μM stock concentration by adding the nmol of primers in the tubes × 10 μL of nuclease-free water to the tube. Vortex briefly and allow to sit at room temperature for 10 minutes before centrifuging to ensure all contents are at their bottoms. Set aside a quantity of 100 μM stock in a -20° freezer to make into 15 μM stocks for the future. Dilute remaining 100 μM stocks in sterile microcentrifuge tubes at a 1:10 ration with nuclease-free water to obtain 10 μM stock for PCR.
2) Turn on Eppendorf Mastercycler Gradient® thermal cycler and turn lid heating on.

3) Because the nosZ and norB primers are run on different PCR cycling programs, as seen in Table 4 and 5, only one primer and its samples can be mixed and run at a time.
4) Thaw all reagents listed in Table 2 and 3 depending on which primer and samples to mix and run first, except GoTaq® DNA Polymerase. It should stay in the freezer until needed in order to protect it from degradation because it does not freeze completely. Thoroughly mix all reagents by light vortexing for a few seconds as various components could have settled out or been distributed unevenly, which will affect the precision of the samples. Lightly centrifuge for a few seconds to ensure all contents are at the bottom before pipetting.
5) The DNA template samples test for nosZ and norB are the pure culture strains Isolate 21 (*Pseudomonas nitroreducers),* PD1222 (*Paracoccus denitrificans)*, Isolate 26 (*Ralstonia pickettii),* Matt's house control: 0 to 1-inch topsoil (MC), and Matt's house leachfield: 0 to 1-inch topsoil (ML). Isolate 21, PD1222, and Isolate 26 were obtained from Shapley lab at Cornell University, grown in Sistrom's Media. Matt's house control and leachfield soil samples were obtained from an Ithacan resident whose septic system is under analysis for this study. DNA was extracted from Isolate 21 and 26 using the MoBio PowerSoil DNA Isolation Kit. DNA was similarly extracted from the soil samples using the aforementioned kit, as well as the MoBio UltraClean Fecal DNA Kit and UltraClean Microbial DNA Isolation Kit.
6) Using calibrated pipetters, create "Master Mixes" for either norB or nosZ according to Table 2 and 3, respectively, in sterile microcentrifuge tubes. Add GoTaq® DNA Polymerase at the very end. Multiply the number of 50 or 30 μL (depending on running norB or nosZ primers) reactions plus two by the amount (μL) per 50 μL or 30 μL to

determine the total amount of Master Mix to make. One additional reaction should be made for the Master Mix because the components are viscous and can be left behind in the pipette tips when mixing. The other additional reaction is for a nuclease-free water sample in place of a DNA template. Any PCR product that appears during gel electrophoresis indicate that the PCR reagents may have been contaminated, thus invalidating the results. Double the amount of Master Mix if wanting to test duplicates of each sample to ensure results are both accurate and precise.

REAGENT STOCK	**AMOUNT (µL) PER 50 µL REACTION**
Promega Corporation 5X Green or Colorless GoTaq® Reaction Buffer	10
Promega Corporation dNTP Mix (10 mM)	1
IDT Forward Primer (10 µM)	5
IDT Reverse Primer (10 µM)	5
Promega Corporation GoTaq® DNA Polymerase (5 Units/µL)	0.2
DNA Template (50 ng) <u>norB DNA Template Samples Tested</u> • Pure Culture Isolate 21 (*Pseudomonas nitroreducers)*: 82.6 ng/µL • Pure Culture PD1222 (*Paracoccus denitrificans)*: 16.6 ng/µL • Matt's house control: 0 to 1 inch topsoil (MC): 9.1 ng/µL • Matt's house leachfield: 0 to 1 inch topsoil (ML): 5.8 ng/µL	Varies depending on pure culture concentration: • Isolate 21: $50\ \text{ng} \times \frac{\mu\text{L}}{82.6\ \text{ng}} = 0.61\ \mu\text{L}$ • PD1222: $50\ \text{ng} \times \frac{\mu\text{L}}{16.6\ \text{ng}} = 3\ \mu\text{L}$ • MC: $50\ \text{ng} \times \frac{\mu\text{L}}{9.1\ \text{ng}} = 5.49\ \mu\text{L}$ • ML: $50\ \text{ng} \times \frac{\mu\text{L}}{16.6\ \text{ng}} = 8.62\ \mu\text{L}$
Promega Corporation Nuclease-Free Water	Varies depending on DNA template amount: • Isolate 21: 28.2 • PD: 25.8 • MC: 23.3 • ML: 20.18

Table 2. NorB Master Mix components and amounts for PCR [19].

REAGENT STOCK	**AMOUNT (μL) PER 30 μL REACTION**
Promega Corporation 5X Green or Colorless GoTaq® Reaction Buffer	6
Promega Corporation dNTP Mix (10 mM)	0.3
IDT Forward Primer (10 μM)	0.6
IDT Reverse Primer (10 μM)	0.6
Promega Corporation GoTaq® DNA Polymerase (5 Units/μL)	0.1
DNA Template (50 ng) <u>nosZ DNA Template Samples Tested</u> • Pure Culture Isolate 26 (*Ralstonia pickettii)*: 16.6 ng/μL • Pure Culture PD1222 (*Paracoccus denitrificans)*: 16.6 ng/μL • Matt's house control: 0 to 1 inch topsoil (MC): 9.1 ng/μL • Matt's house leachfield: 0 to 1 inch topsoil (ML): 5.8 ng/μL	Varies depending on pure culture concentration: • Isolate 26: $50\ \text{ng} \times \frac{\mu\text{L}}{16.6\ \text{ng}} = 3\ \mu\text{L}$ • PD1222: $50\ \text{ng} \times \frac{\mu\text{L}}{16.6\ \text{ng}} = 3\ \mu\text{L}$ • MC: $50\ \text{ng} \times \frac{\mu\text{L}}{9.1\ \text{ng}} = 5.49\ \mu\text{L}$ • ML: $50\ \text{ng} \times \frac{\mu\text{L}}{16.6\ \text{ng}} = 8.62\ \mu\text{L}$
Promega Corporation Nuclease-Free Water	Varies depending on DNA template amount: • Isolate 21: 25.8 • PD: 25.8 • MC: 23.3 • ML: 20.18

Table 3. NosZ Master Mix components and amounts for PCR [11].

7) Cap Master Mix tube. Mix the Master Mix well by inverting or by brief, low-setting vortexing. Vortexing too long or strongly can damage the GoTaq® DNA Polymerase.
8) For norB, distribute 49.39, 47, 44.51, and 41.38 μL of Master Mix into 4 labeled 0.2 μL sterile PCR tubes according to calculations done in Table 2. Add 47 μL of Master Mix into another labeled 0.2 μL sterile PCR tube for the nuclease-free water sample, bearing in mind that because this is the sterile sample, any reasonable amount of Master Mix can be used. Replicate this if making duplicates. For nosZ, distribute 47, 47, 44.51, and 41.38

μL of Master Mix into 4 labeled 0.2 μL sterile PCR tubes according to calculations done in Table 3. Again, replicate if making duplicates and add 47 μL of Master Mix into another 0.2 labeled μL sterile PCR tube for the sterile sample.

9) For norB, add 0.61, 3, 5.49, 8.62 μL of corresponding DNA template samples into the corresponding 0.2 μL PCR tubes according to calculations done in Table 2. For the sterile control, use 3 μL of nuclease-free water, bearing in mind that any reasonable amount of sterile water can be used. For norB, add 3, 3, 5.49, 8.62 μL of corresponding DNA template samples into the corresponding 0.2 μL PCR tubes according to calculations done in Table 3. For the sterile control, use 3 μL of nuclease-free water.
10) Cap the PCR tubes. Flick to mix, invert, and vortex briefly. Ensure that the contents are at the bottom of the tubes. Tapping the tubes on a lab surface or a brief three-second pulse in a microcentrifuge will send the samples to the bottoms of the tubes.
11) Place tubes into Eppendorf Mastercycler Gradient® thermal cycler. Program the thermal cycler to run the thermal cycling programs according to Table 4 and 5, depending on whether testing nosZ or norB primers first.

CYCLE STEP	TEMPERATURE (°C)	TIME (min:sec)	CYCLES
Initial Denaturation	95	10:00	1
Denaturation	95	00:30	30
Annealing	60	00:45	30
Extension	72	00:45	30
Final Extension	72	05:00	1
HOLD	4	HOLD	HOLD

Table 4. NorB PCR thermal cycling program [18].

CYCLE STEP	TEMPERATURE (°C)	TIME (min)	CYCLES
Initial Denaturation	95	10:00	1
Denaturation	95	00:30	30
Annealing	52	00:45	30
Extension	72	00:45	30
Final Extension	72	05:00	1
HOLD	4	HOLD	HOLD

Table 5. NosZ PCR thermal cycling program [18].

12) After the samples are finished running, remove from thermal cycler and store in 20° freezer until testing with gel electrophoresis. If running gel electrophoresis immediately after, cool and keep samples on ice to prevent degradation.

GEL ELECTROPHORESIS

1) Prepare 0.5X TBE Buffer for use in agarose gel electrophoresis. To make 5X TBE Stock Solution Buffer, in a 2-liter flask, mix 108 grams of Tris base, 55 grams of Boric acid, 40 mL 0.5 M Ethylenediaminetetraacetic acid (EDTA) pH 8.0, and 1600 mL distilled water. Add a large magnetic stir bar and place on stir plate on high to mix. Low heat can be applied to facilitate dissolving. Once dissolved, add water to adjust volume to 2 liters. Precipitation will form gradually but the buffer is still usable upon further mixing and low heating. To prepare a working solution of 0.5X TBE Buffer, add 100 mL of 5X TBE Buffer to 900 mL of distilled deionized water and mix [22].
2) In an Erlenmeyer flask, prepare gel solution by combining required volume of 0.5X TBE Buffer with 1% (wt/vol) agarose. The total volume needed is 150 mL and 1.5 grams of agarose for a large gel or 50 mL and 0.5 grams of agarose for a small gel.

3) To dissolve and boil solution, place the flask with loosely stuffed Kimwipes in the neck into the microwave. Microwave for 30 seconds. Remove, swirl gently, and microwave an additional 30 seconds or until it just starts to boil.
4) Remove and cool on the bench for about 15 minutes until it is cool enough to pour into the gel mold. Dilute the Biotium GelRed™10,000X stock reagent into the molten agarose gel solution at 1:10,000 and mix thoroughly [23]. A large gel would require 15 μL of GelRed™.
5) Pour the gel solution into the gel mold positioned on flat surface. Check beforehand with the small level under the gel box. Immediately place the well-comb in the grooves in the gel box. The comb(s) suspended in the gel create wells to load the PCR products into. Allow to solidify at least 20 minutes.
6) When the gel is solidified, move it to the gel box. Fill the gel box with the same 0.5X TBE Buffer used to make the gel until it covers the gel by a few millimeters. Remove the comb(s) gently.
7) Thaw and keep PCR products on ice. Thaw Thermo Scientific 100 base pair DNA Ladder for testing PCR products sizes to a standard. Thaw Thermo Scientific 6X DNA Gel Loading Dye. This aids in loading the samples into the gel and in checking on the extent of gel electrophoresis over time. Thoroughly mix all reagents during the set-up procedure by light vortexing for a few seconds as various components could have settled out or been distributed unevenly, which will affect the precision of the samples. Lightly centrifuge for a few seconds to ensure all contents are at the bottom before pipetting.

8) Cut a piece of Parafilm and place a 1 μL spot of loading dye on it for each sample and standard. For each standard or sample, draw up 3-5 μL of PCR product and combine with the 1 μL of loading dye by pipetting up and down gently. Pipette up as much volume as possible but avoid any air bubbles as these make it tough to load the gel effectively.

9) Holding the pipette in the middle of a fresh gel well, dispense the sample/dye mix. Do not puncture the bottom of the well with the pipette tip. Also, do not push the pipette to the second "stop" as this will introduce an air bubble which will rise up out of the well and carry most of the sample with it.

10) When all samples have been loaded, place the cover on the gel box and ensure that the electrical leads are plugged into the Bio-Rad PowerPac 300 Electrophoresis Power Supply. Turn on the power supply to provide 100 Volts of potential across the leads. Hit "RUN" and watch for small bubbles rising in the gel box at the two electrodes. Allow to run for at least 45 minutes, checking to see where the dye molecules are. After the desired amount of time, around 1.5 hours, hit "STOP" to stop the electrophoresis process. Remove the cover of the gel box.

11) Using a large spatula, move the gel onto the surface of the Fisher Scientific UVP Benchtop UV Transilluminators: Single UV, which provides UV illumination of the stained DNA. The GelRed™ will have caused the bands to turn red. The bands of the PCR products should correspond to the size of the bands of the DNA ladder.

12) Discard the gel in the trash after taking an image. If desired, store in the gel at 4° C in a Ziploc bag for future use.

DESIGNING LONG AMPLICON PRIMERS

Long amplicon primers for nosZ and norB must be designed and used to amplify the degenerate primer amplicon targeted sequences present in the small and limited quantities of soil samples obtained from each testing site. The ideal placement of the degenerate primers relative to the long amplicon primers is shown in Figure 4. This is done in order to create a large volume of PCR products from the degenerate primers for the analysis and the creation of standards, without needing to use the entirety of the original soil samples.

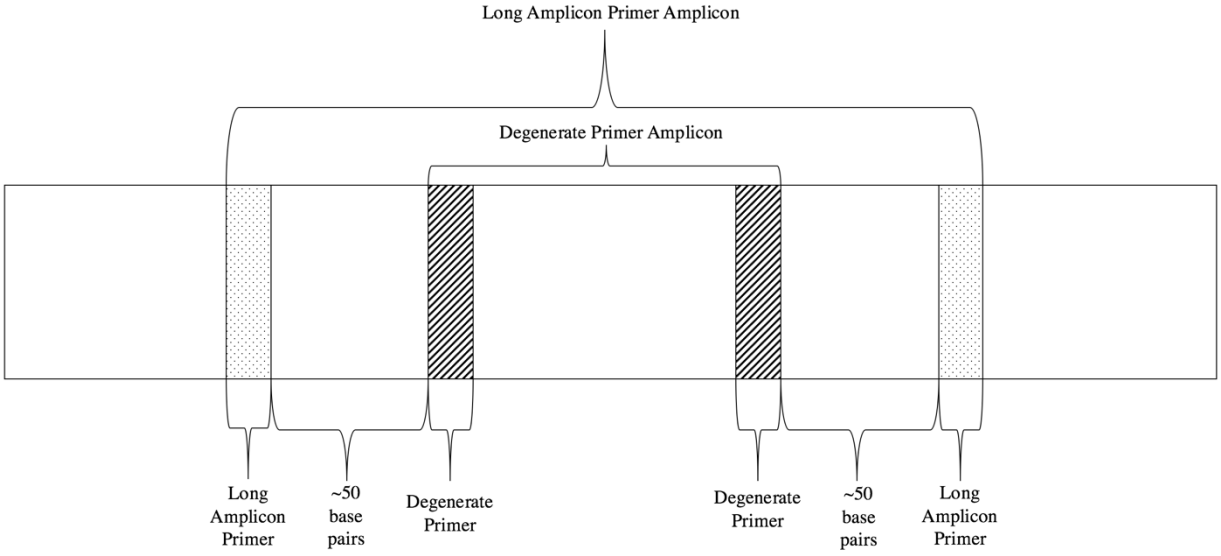

Figure 4. Diagram of Degenerate and Long Amplicon Primers Placement in DNA Sequence.

JGI IMG/M PROTOCOL

The Department of Energy's (DOE) Joint Genome Institute (JGI) is an international, free database which enables the annotation, analysis, and distribution of sequenced microbial genome and microbiome datasets. JGI contains the system Integrated Microbial Genomes & Microbiomes (IMG/M), which is where the gene sequences of nosZ and norB from *Paracoccus*

denitrificans PD1222 can be obtained [24]. Steps for obtaining the gene sequences are as follows:

1) On the front page, click on *Find Genomes* on the top bar and then *Genome Search*. Type in the keyword *Paracoccus denitrificans* PD1222 with the filter *Genome Name.*
2) Select the genome with the IMG Genome ID of 639633048 and click *Add Selected to Genome Cart* so that the gene sequences for nosZ and norB can be selected only from *Paracoccus denitrificans* PD1222.
3) On the top bar, click on *Find Genes* and then *Gene Search.* Because the filter chosen is *Gene Product Name (inexact),* under keyword, type in nitrous oxide reductase if looking for nosZ and nitric oxide reductase if looking for norB.
4) Under *Sequencing Status*, *Paracoccus denitrificans* PD1222 should still be in Genome Cart. Add it to *Selected Genomes*. Click *Go.* When the search results for gene product name appear for nitric oxide reductase, select the gene product name for NorB subunit apoprotein with Gene ID 639769812. The Gene ID for nitrous oxide reductase should be 639771526.
5) *Gene Information* should now appear. To view gene sequences of nosZ or norB, click on the link next to *DNA Coordinates*. Copy sequences as shown in Figure 5 and 6:

>639771526 nitrous oxide reductase apoprotein [Paracoccus denitrificans PD1222 chromosome 2: NC_008687] (-) strand

ATGGAATCGAAACAGGAAAAGGGTCTCAGCCGCCGCGCGCTTTTGGGCGCGACAGCCGGGGGCGCGGCCGTGGCCGG
CGCGTTCGGCGGACGGCTGGCGCTGGGGCCGGCGGCGCTCGGCCTGGGCACGGCGGGGGTCGCGACCGTCGCCGGCA
GTGGTGCGGCGCTGGCCGCCAGCGGCGACGGCTCGGTCGCGCCGGGGCAACTGGACGACTACTACGGCTTCTGGTCC
TCGGGCCAGAGCGGCGAGATGCGCATCCTTGGCATTCCCTCGATGCGCGAGCTGATGCGGGTTCCGGTCTTCAACCG
CTGCTCGGCCACCGGCTGGGGCCAGACCAACGAATCGCTGCGCATCCATGAGCGCACCATGAGCGAGCGGACGAAGA
AGTTCCTTGCCGCCAATGGCAAGCGCATCCACGACAACGGCGACCTGCACCACGTCCACATGTCCTTTACCGAGGGC
AAATATGACGGCCGCTTCCTGTTCATGAACGACAAGGCCAATACCCGCGTGGCGCGGGTGCGCTGCGACGTGATGAA
ATGCGACGCCATCCTGGAGATCCCCAACGCCAAGGGCATCCACGGCCTGCGGCCGCAGAAATGGCCGCGCAGCAACT
ATGTCTTCTGCAATGGCGAGGACGAGACGCCGCTGGTCAACGACGGCACCAACATGGAGGACGTGGCGAATTACGTG
AACGTCTTCACCGCCGTCGATGCCGACAAGTGGGAGGTCGCCTGGCAGGTGCTCGTCTCGGGCAACCTGGACAACTG
CGATGCCGATTACGAGGGCAAATGGGCCTTTTCGACCTCCTACAACTCGGAAAAGGGGATGACCCTGCCCGAGATGA
CGGCGGCCGAGATGGACCATATCGTCGTCTTCAACATCGCCGAGATCGAAAAGGCCATCGCTGCCGGCGACTATCAG
GAACTGAACGGCGTCAAGGTCGTGGATGGCCGGAAAGAGGCAAGCAGCCTGTTCACCCGCTATATCCCCATCGCCAA
CAACCCGCATGGCTGCAACATGGCGCCGGACAAGAAGCACCTGTGCGTCGCGGGCAAGCTGTCGCCCACGGTGACGG
TGCTGGACGTGACCCGGTTCGACGCGGTGTTCTATGAGAACGCCGATCCGCGCAGCGCCGTGGTGGCCGAGCCGGAA
CTGGGCCTTGGCCCGCTGCACACCGCCTTCGACGGGCGCGGCAACGCCTATACC**TCGCTGTTCCTCGACAGCCAGGT**
GGTCAAGTGGAACATCGAGGATGCCATCCGCGCCTATGCCGGCGAGAAGGTCGACCCGATCAAGGACAAGCTGGACG
TGCATTACCAGCCCGGCCACCTCAAGACCGTGATGGGCGAGACGCTGGACGCCACCAACGACTGGCTGGTCTGCCTG
TCCAAGTTCTCGAAGGACCGCTTCCTGAACGTCGGGCCGCTGAAGCCGGAGAACGACCAGCTGATCGACAT**CTCGGG**
CGACAAGATGGTGCTGGTCCATGACGGCCCCACCTTTGCCGAGCCGCACGACGCCATCGCCGTGCACCCCTCGATCC
TGTCCGACATCAAGTCGGTCTGGGACCGCAACGACCCGATGTGGGCCGAGACCCGCGCTCAGGCCGAGGCCGATGGC
GTCGACATCGACAACTGGACCGAAGAGGTGATCCGCGACGGCAACAAGGTGCGGGTCTACATGTCCAGCGTGGCGCC
GAGCTTCTCGATCGAGAGCTTTACCGTCAAGGAGGGCGACGAGGTCACGGTGATCGTCACCAACCTGGACGAGATCG
ACGATCTGACGCATGGCTTCACCATGGGCAACTACGGCGTCGCCATGGAGATCGGGCCGCAGATGACCAGCTCGGTC
ACCTTCGTCGCCGCCAATCCCGGGGTCTATTGGTATTATTGCCAATGGTTCTGCCATGCGCTGCACATGGAGATGCG
CGGCCGCATGCTGGTCGAGCCGAAGGAGGCCTGA

Figure 5. NosZ nucleic acid base sequence with degenerate nosZ forward and reverse primers in bold with light gray background. Degenerate primer amplicon sequence amplified is in bold lettering.

>639769812 nitric oxide reductase, NorB subunit apoprotein [Paracoccus denitrificans PD1222 chromosome 1: NC_008686] (-) strand

ATGAGATATCATTCGCAACGCATCGCCTATGCCTATTTCCTTGTGGCCATGGTGCTTTTCGCCGTGCAGGTCACGAT
CGGCCTGATCATGGGCTGGATCTATGTCAGCCCGAACTTCCTGTCCGAACTCCTGCCCTTCAACATCGCCCGGATGC
TGCACACCAACAGCCTGGTCGTCTGGCTGCTCCTGGGCTTTTTCGGCGCGACCTATTACATCCTGCCCGAAGAGGCC
GAGCGCGAGATCCATTCGCCGCTGCTGGCCTGGATCCAGCTGGGCATCTTCGTGCTTGGCACGGCGGGCGTGGTCGT
GACCTATCTGTTCGACCTGTTCCACGGCCACTGGCTGCTGGGCAAGGAGGGGCGCGAGTTCCTGGAACAGCCGAAAT
GGGTCAAGCTCGGCATCGCCGTCGCCGCGGTGATCTTCATGTACAACGTCAGCATGACCGCGCTGAAGGGGCGGCGG
ACGGCGGTGACCAACGTGCTGCTGATGGGCCTTTGGGGCCTAGTGCTGCTGTGGCTCTTTGCCTTCTACAACCCGGC
GAACCTGGTTCTC**GACAAGCAATACTGGT**GGTGGGTCATCCATCTGTGGGTCGAGGGCGTGTGGGAGCTGATCATGG
CAGCCATCCTTGCCTTCCTGATGCTCAAGCTGACCGGCGTGGACCGCGAGGTGGTCGAGAAATGGCTTTACGTCATC
GTCGCCACGGCGCTGTTCTCGGGCATCCTGGGCACCGGGCACCATTACTACTGGATCGGCCTGCCGGCTTACTGGCA
GTGGATCGGCTCGATCTTCTCGAGCTTCGAGATCGTGCCCTTCTTCGCCATGATGTCATTCGCCTTCGTCATGGTCT
GGAAGGGCCGGCGCGACCATCCGAACAAGGCCGCGCTGGTCTGGAGCCTGGGCTGCACCGTGCTGGCCTTCTTCGGC
GCTGGGGTCTGGGGCTTCCTGCACACGCTGCACGGGGTGAACTACTACACCCACGGCACGCAGATCACCGCCGCGCA
CGGCCACCTGGCCTTCTATGGCGCCTATGTCTGCCTGGTGCTGGCGCTGGTGACCTATTGCATGCCGCTGATGAAGA
ACCGCGACCCCTACAACCAGGTGCTGAACATGGCCTCGTTCTGGCTGATGTCCTCGGGCATGGTGTTCATGACGGTG
ACGCTGACCTTCGCCGGCACGGTGCAGACCCATCTGCAACGCGTCGAGGGCGGGTTCTTCATGGACGTGCAGGACGG
GCTGGCGCTGTTCTACTGGATGCGCTTCGGTTCGGGCGTGGCCGTCGTGCTGGGGGCGCTGCTGTTCATCTATGCCG
TGCTGTTCCCACGCCGCGAGGTGGTCAAGGCCGGTCCGGTGCAGGCGCACAAGGACGGTCATCTGGAGGCTGCGGAG
TAA

Figure 6. NorB nucleic acid base sequence with degenerate norB forward and reverse primers in bold with light gray background. Degenerate primer amplicon sequence amplified is in bold lettering.

IDT PROTOCOL

The long amplicon primers were designed and ordered from Integrated DNA Technologies (IDT), which supplies custom nucleic acids and design tools [25]. Steps for designing and ordering both the norB and nosZ primers are as follows:

1) Under their *Products & Services*, there is a link to *PrimerQuest Tool* underneath the category Gene Expression.
2) Beneath *Enter sequence(s) manually*, paste the norB or nosZ nucleic acid base sequences found from IMG/M into the Textbox. Alternatively, the sequences(s) can be downloaded using Genbank or Accession ID, or uploaded via an Excel file.
3) Click *Show Custom Design Parameters*. Because the sequences for the qPCR primers of nosZ and norB must be included in the final amplicon product which the long amplicon primers generate, the *Excluded Region* should have 1210 base pairs-1457 base pairs for nosZ primers and 553 base pairs-942 base pairs for norB primers.
4) Because long amplicon primers should be approximately 50 base pairs before and after the start and end of the qPCR primer regions, to focus the location of the suggested long amplicon primers to this area, the *Included Region* can include about 100 base pairs before and after the region. For example, the start-end for nosZ could be 1110-1557, while the start-end for norB could be 453-1042.
5) Under *Primer Details,* ensure that the primer length is somewhere between 17-25 nucleotides (nt). The primer Tm (°C), or the melting temperature, can be adjusted to fit the an existing PCR or qPCR protocol, but should have a Tm of 50-70°, ideally between 55-65°. To estimate whether the primers designed have this ideal Tm range, employ the McConaughy equation and Wallace equation [26].

a. McConaughy Equation (Biochemistry 8: 3289-3295, 1969) modified for cycle sequencing: Tm = 60 + 41(G + C)/L -500/L where L = length of primer.
b. Wallace Equation (Nucleic Acids Research 6: 3543-3557, 1979): Dissociation temperature (Td) = 2(A + T) + 4(G + C)

The primer GC (%) content should be approximately 50% of the total length. In addition, the Tm of the 5' end of the primers should be similar to the Tm of the 3' end, which is determined by counting the number of A/T bases and C/G bases within 6 nucleotides of each end. Choosing the primers with the most similar numbers ensures that the primers anneal flat with the template strand. In addition, the avoid primers that can form hairpin loops or primer dimers, it is best to avoid stretches of more than 2 identical bases, especially C or G, at the 3' end. It may cause mismatches or slippage during annealing, resulting a bulge of the primer/template hybrid which prevents the DNA polymerase from attaching [26].

6) After the *Custom Design Parameters* have been filled in, click *Get Assays.* A selection of suggested sequence results will appear. Click on the link *View Assay Details* to verify that the primers are at the desired location.
7) If satisfied, select the sequence desired and click *Add to Order.* Under *Order Options*, select *Forward & Reverse*, as in forward and reverse primers, and then click *Oligos in Tubes*, which means the gene sequences will come in desiccated form to be rehydrated.
8) Name the sequences selected to be one that is easily recognizable. The *Scale Options* should be 25 nmole DNA oligo. Select *Add to Cart.*

9) Before selecting *Checkout*, the contents of the cart including the quote can be emailed before ordering or added to a wish list.

10) The long amplicon primers will arrive in 2-3 days in a desiccated form to be rehydrated for testing. The placement of the long amplicon primers in conjunction with the degenerate primers on the nosZ and norB stretch on Paracoccus denitrificans PD1222 are shown in Figure 7, Figure 8, and Figure 9.

>639771526 nitrous oxide reductase apoprotein [Paracoccus denitrificans PD1222 chromosome 2: NC_008687] (-) strand

ATGGAATCGAAACAGGAAAAGGGTCTCAGCCGCCGCGCGCTTTTGGGCGCGACAGCCGGGGGCGCGGCCGTGGCCGG
CGCGTTCGGCGGACGGCTGGCGCTGGGGCCGGCGGCGCTCGGCCTGGGCACGGCGGGGGTCGCGACCGTCGCCGGCA
GTGGTGCGGCGCTGGCCGCCAGCGGCGACGGCTCGGTCGCGCCGGGGCAACTGGACGACTACTACGGCTTCTGGTCC
TCGGGCCAGAGCGGCGAGATGCGCATCCTTGGCATTCCCTCGATGCGCGAGCTGATGCGGGTTCCGGTCTTCAACCG
CTGCTCGGCCACCGGCTGGGGCCAGACCAACGAATCGCTGCGCATCCATGAGCGCACCATGAGCGAGCGGACGAAGA
AGTTCCTTGCCGCCAATGGCAAGCGCATCCACGACAACGGCGACCTGCACCACGTCCACATGTCCTTTACCGAGGGC
AAATATGACGGCCGCTTCCTGTTCATGAACGACAAGGCCAATACCCGCGTGGCGCGGGTGCGCTGCGACGTGATGAA
ATGCGACGCCATCCTGGAGATCCCCAACGCCAAGGGCATCCACGGCCTGCGGCCGCAGAAATGGCCGCGCAGCAACT
ATGTCTTCTGCAATGGCGAGGACGAGACGCCGCTGGTCAACGACGGCACCAACATGGAGGACGTGGCGAATTACGTG
AACGTCTTCACCGCCGTCGATGCCGACAAGTGGGAGGTCGCCTGGCAGGTGCTCGTCTCGGGCAACCTGGACAACTG
CGATGCCGATTACGAGGGCAAATGGGCCTTTTCGACCTCCTACAACTCGGAAAAGGGGATGACCCTGCCCGAGATGA
CGGCGGCCGAGATGGACCATATCGTCGTCTTCAACATCGCCGAGATCGAAAAGGCCATCGCTGCCGGCGACTATCAG
GAACTGAACGGCGTCAAGGTCGTGGATGGCCGGAAAGAGGCAAGCAGCCTGTTCACCCGCTATATCCCCATCGCCAA
CAACCCGCATGGCTGCAACATGGCGCCGGACAAGAAGCACCTGTGCGTCGCGGGCAAGCTGTCGCCCACGGTGACGG
TGCTGGACGTGACCCGGTTCGACGCGGTGTTCTATGAGAACGCCGATCCGCGCAGCGCCGTGGTGGCCGAGCCGGAA
CTGGGCCTTGGCCCGCTGCACACCGCCTTCGACGGGCGCGGCAACGCCTATACC**TCGCTGTTCCTCGACAGCCAGGT**
GGTCAAGTGGAACATCGAGGATGCCATCCGCGCCTATGCCGGCGAGAAGGTCGACCCGATCAAGGACAAGCTGGACG
TGCATTACCAGCCCGGCCACCTCAAGACCGTGATGGGCGAGACGCTGGACGCCACCAACGACTGGCTGGTCTGCCTG
TCCAAGTTCTCGAAGGACCGCTTCCTGAACGTCGGGCCGCTGAAGCCGGAGAACGACCAGCTGATCGACATCTCGGG
CGACAAGATGGTGCTGGTCCATGACGGCCCCACCTTTGCCGAGCCGCACGACGCCATCGCCGTGCACCCCTCGATCC
TGTCCGACATCAAGTCGGTCTGGGACCGCAACGACCCGATGTGGGCCGAGACCCGCGCTCAGGCCGAGGCCGATGGC
GTCGACATCGACAACTGGACCGAAGAGGTGATCCGCGACGGCAACAAGGTGCGGGTCTACATGTCCAGCGTGGCGCC
GAGCTTCTCGATCGAGAGCTTTACCGTCAAGGAGGGCGACGAGGTCACGGTGATCGTCACCAACCTGGACGAGATCG
ACGATCTGACGCATGGCTTCACCATGGGCAACTACGGCGTCGCCATGGAGATCGGGCCGCAGATGACCAGCTCGGTC
ACCTTCGTCGCCGCCAATCCCGGGGTCTATTGGTATTATTGCCAATGGTTCTGCCATGCGCTGCACATGGAGATGCG
CGGCCGCATGCTGGTCGAGCCGAAGGAGGCCTGA

Figure 7. NosZ nucleic acid base sequence with 1st set of nosZ long amplicon forward and reverse primers underlined. Degenerate nosZ forward and reverse primers are bolded with gray background. Degenerate primers amplicon sequence amplified is in bold.

>639771526 nitrous oxide reductase apoprotein [Paracoccus denitrificans PD1222 chromosome 2: NC_008687] (-) strand

ATGGAATCGAAACAGGAAAAGGGTCTCAGCCGCCGCGCGCTTTTGGGCGCGACAGCCGGGGGCGCGGCCGTGGCCGG
CGCGTTCGGCGGACGGCTGGCGCTGGGGCCGGCGGCGCTCGGCCTGGGCACGGCGGGGGTCGCGACCGTCGCCGGCA
GTGGTGCGGCGCTGGCCGCCAGCGGCGACGGCTCGGTCGCGCCGGGGCAACTGGACGACTACTACGGCTTCTGGTCC
TCGGGCCAGAGCGGCGAGATGCGCATCCTTGGCATTCCCTCGATGCGCGAGCTGATGCGGGTTCCGGTCTTCAACCG
CTGCTCGGCCACCGGCTGGGGCCAGACCAACGAATCGCTGCGCATCCATGAGCGCACCATGAGCGAGCGGACGAAGA

AGTTCCTTGCCGCCAATGGCAAGCGCATCCACGACAACGGCGACCTGCACCACGTCCACATGTCCTTTACCGAGGGC
AAATATGACGGCCGCTTCCTGTTCATGAACGACAAGGCCAATACCCGCGTGGCGCGGGTGCGCTGCGACGTGATGAA
ATGCGACGCCATCCTGGAGATCCCCAACGCCAAGGGCATCCACGGCCTGCGGCCGCAGAAATGGCCGCGCAGCAACT
ATGTCTTCTGCAATGGCGAGGACGAGACGCCGCTGGTCAACGACGGCACCAACATGGAGGACGTGGCGAATTACGTG
AACGTCTTCACCGCCGTCGATGCCGACAAGTGGGAGGTCGCCTGGCAGGTGCTCGTCTCGGGCAACCTGGACAACTG
CGATGCCGATTACGAGGGCAAATGGGCCTTTTCGACCTCCTACAACTCGGAAAAGGGGATGACCCTGCCCGAGATGA
CGGCGGCCGAGATGGACCATATCGTCGTCTTCAACATCGCCGAGATCGAAAAGGCCATCGCTGCCGGCGACTATCAG
*GAACTGAACGGCGTCAAGGTCGTGGATGGCCGGAAAGAGGCAAGCAG**CCTGTTCACCCGCTATATCC**CCATCGCCAA*
CAACCCGCATGGCTGCAACATGGCGCCGGACAAGAAGCACCTGTGCGTCGCGGGCAAGCTGTCGCCCACGGTGACGG
TGCTGGACGTGACCCGGTTCGACGCGGTGTTCTATGAGAACGCCGATCCGCGCAGCGCCGTGGTGGCCGAGCCGGAA
*CTGGGCCTTGGCCCGCTGCACACCGCCTTCGACGGGCGCGGCAACGCCTATACC**TCGCTGTTCCTCGACAGCCAGGT***
GGTCAAGTGGAACATCGAGGATGCCATCCGCGCCTATGCCGGCGAGAAGGTCGACCCGATCAAGGACAAGCTGGACG
TGCATTACCAGCCCGGCCACCTCAAGACCGTGATGGGCGAGACGCTGGACGCCACCAACGACTGGCTGGTCTGCCTG
***TCCAAGTTCTCGAAGGACCGCTTCCTGAACGTCGGGCCGCTGAAGCCG**GAGAACGACCAGCTGATCGACAT**CTCGGG*
CGACAAGATGGTGCTGGTCCATGACGGCCCCACCTTTGCCGAGCCGCACGACGCCATCGCCGTGCACCCCTCGATCC
TGTCCGACATCAAGTCGGTCTGGGACCGCAACGACCCGATGTGGGCCGAGACCCGCGCTCAGGCCGAGGCCGATGGC
*GTCGACATCGACAACTGGACCGAAGAGGTGA**TCCGCGACGGC**AACAAGGTGCGGGTCTACATGTCCAGCGTGGCGCC*
GAGCTTCTCGATCGAGAGCTTTACCGTCAAGGAGGGCGACGAGGTCACGGTGATCGTCACCAACCTGGACGAGATCG
ACGATCTGACGCATGGCTTCACCATGGGCAACTACGGCGTCGCCATGGAGATCGGGCCGCAGATGACCAGCTCGGTC
ACCTTCGTCGCCGCCAATCCCGGGGTCTATTGGTATTATTGCCAATGGTTCTGCCATGCGCTGCACATGGAGATGCG
CGGCCGCATGCTGGTCGAGCCGAAGGAGGCCTGA

Figure 8. NosZ nucleic acid base sequence with 2nd set of nosZ long amplicon forward and reverse primers bolded, shaded, and underlined. 1st set of nosZ long amplicon forward and reverse primers is underlined. Degenerate nosZ forward and reverse primers are bolded with gray background. Degenerate primers amplicon sequence amplified is in bold.

>639769812 nitric oxide reductase, NorB subunit apoprotein [Paracoccus denitrificans PD1222 chromosome 1: NC_008686] (-) strand

ATGAGATATCATTCGCAACGCATCGCCTATGCCTATTTCCTTGTGGCCATGGTGCTTTTCGCCGTGCAGGTCACGAT
CGGCCTGATCATGGGCTGGATCTATGTCAGCCCGAACTTCCTGTCCGAACTCCTGCCCTTCAACATCGCCCGGATGC
TGCACACCAACAGCCTGGTCGTCTGGCTGCTCCTGGGCTTTTTCGGCGCGACCTATTACATCCTGCCCGAAGAGGCC
GAGCGCGAGATCCATTCGCCGCTGCTGGCCTGGATCCAGCTGGGCATCTTCGTGCTTGGCACGGCGGGCGTGGTCGT
GACCTATCTGTTCGACCTGTTCCACGGCCACTGGCTGCTGGGCAAGGAGGGGCGCGAGTTCCTGGAACAGCCGAAAT
GGGTCAAGCTCGGCATCGCCGTCGCCGCGGTGATCTTCATGTACAACGTCAGCATGACCGCGCTGAAGGGGCGGCGG
ACGGCGGTGACCAACGTGCTGCTGATGGGCCTTTGGGGCCTAGTGCTGCTGTGGCTCTTTGCCTTCTACAACCCGGC
GAACCTGGTTCTC**GACAAGCAATACTGGTGGTGGGTCATCCATCTGTGGGTCGAGGGCGTGTGGGAGCTGATCATGG**
CAGCCATCCTTGCCTTCCTGATGCTCAAGCTGACCGGCGTGGACCGCGAGGTGGTCGAGAAATGGCTTTACGTCATC
GTCGCCACGGCGCTGTTCTCGGGCATCCTGGGCACCGGGCACCATTACTACTGGATCGGCCTGCCGGCTTACTGGCA
GTGGATCGGCTCGATCTTCTCGAGCTTCGAGATCGTGCCCTTCTTCGCCATGATGTCATTCGCCTTCGTCATGGTCT
GGAAGGGCCGGCGCGACCATCCGAACAAGGCCGCGCTGGTCTGGAGCCTGGGCTGCACCGTGCTGGCCTTCTTCGGC
GCTGGGGTCTGGGGCTTCCTGCACACGCTGCACGGGGTGAACTACTACACCCACGGCACGCAGATCACCGCCGCGCA
CGGCCACCTGGCCTTCTATGGCGCCTATGTCTGCCTGGTGCTGGCGCTGGTGACCTATTGCATGCCGCTGATGAAGA
ACCGCGACCCCTACAACCAGGTGCTGAACATGGCCTCGTTCTGGCTGATGTCCTCGGGCATGGTGTTCATGACGGTG
ACGCTGACCTTCGCCGGCACGGTGCAGACCCATCTGCAACGCGTCGAGGGCGGGTTCTTCATGGACGTGCAGGACGG
GCTGGCGCTGTTCTACTGGATGCGCTTCGGTTCGGGCGTGGCCGTCGTGCTGGGGGCGCTGCTGTTCATCTATGCCG
TGCTGTTCCCACGCCGCGAGGTGGTCAAGGCCGGTCCGGTGCAGGCGCACAAGGACGGTCATCTGGAGGCTGCGGAG
TAA

Figure 9. NorB nucleic acid base sequence with norB long amplicon forward and reverse primers underlined. Degenerate norB forward and reverse primers are bolded with gray background. Degenerate primers amplicon sequence amplified is in bold.

VALIDATING LONG AMPLICON PRIMERS

Verifying the long amplicon primers designed and shown in Table 6 through end-point PCR and gel electrophoresis must be done in order to ensure the effectiveness of their specific binding and amplification of their PCR products. In addition, high throughout sequencing must be conducted to verify that the DNA sequence amplified is indeed the sizably larger long amplicon primer amplicon with the degenerate primer amplicon as well.

PRIMER	**5' TO 3' SEQUENCE**	**AMPLICON SIZE (base pairs)**
nosZ long amp F 1st set	GATGGACCATATCGTCGTCTTC	791
nosZ long amp R 1st set	TCACCTCTTCGGTCCAGTT	791
nosZ long amp F 2nd set	CCTGTTCACCCGCTATATCC	706
nosZ long amp R 2nd set	GTAGACCCGCACCTTGTT	706
norB long amp F	TGCTGATGGGCCTTTGG	542
norB long amp R	GCCATAGAAGGCCAGGTG	542

Table 6. Long Amplicon denitrification pathway primers.

END-POINT PCR

The protocol for end-point PCR is approximately the same as the protocol for validating the degenerate primers, with a few exceptions. Because these long amplicon primers have not been tested in literature for their ideal PCR thermal cycling temperatures, a range of annealing temperatures must be tested to ascertain their optimal specificity:

1) The Eppendorf Mastercycler Gradient® thermal cycler allows one to choose a gradient of temperatures for the annealing step, with the temperatures steadily increasingly horizontally. Thus, the thermal cycling program as seen in Table 7 will have its annealing temperatures adjusted to fit a gradient of temperatures instead of just one set temperature as before. Both primers sets can be run on one plate.

CYCLE STEP	TEMPERATURE (°C)	TIME (min:sec)	CYCLES
Initial Denaturation	95	10:00	1
Denaturation	95	00:30	30
Annealing	53-60	00:45	30
Extension	72	00:45	30
Final Extension	72	05:00	1
HOLD	4	HOLD	HOLD

Table 7. Long amplicon norB and nosZ primer PCR thermal cycling program [18].

2) Because the plate of the PCR machine can hold up to 96 reactions, the PCR reactions to be tested are the 5 DNA templates (Isolate 21, PD1222, MC, ML, and Isolate 26), 8 annealing temperatures (53-60°C), and 2 long amplicon primer sets (norB and nosZ) for a total of 64 reactions, as demonstrated in Figure 10 below.

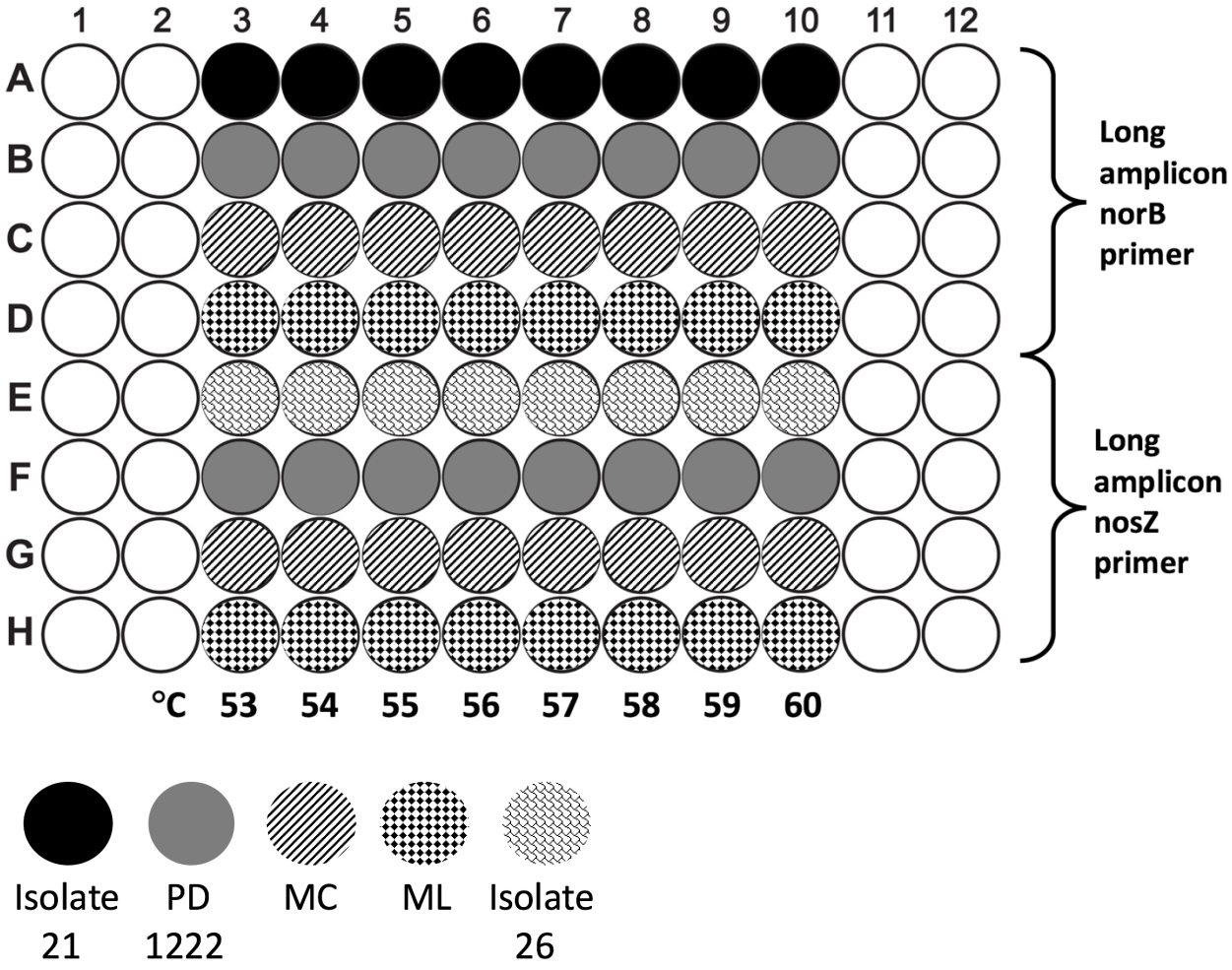

Figure 10: End-point PCR Layout to find optimal annealing temperature.

GEL ELECTROPHORESIS

The protocol for gel electrophoresis should be the same as the ones used for validating the PCR primers. It is recommended in addition to the Thermo Scientific 100 base pair DNA Ladder, to also use the Thermo Scientific Low Range DNA Ladder to more easily identify the amplicon sizes of the long amplicon primers, which are now several hundred base pairs longer than those of the degenerate primers. In addition, after examining the intensity of the bands generated from the reactions with different annealing temperatures, several of the strongest with the highest specificity should be prepared for high-throughput sequencing, which will verify via quantitative measurement the concentration of PCR product generated, not just a qualitative measurement.

PREPARATION FOR HIGH-THROUGHPUT SEQUENCING

Measuring DNA Concentration with NanoDrop

The NanoDrop 2000c measures the concentration of DNA found in the sample. It provides a quantitative measurement of the PCR products generated, so that before continuing with the rest of the preparation for high-throughput sequencing, one can assess whether there is sufficient purity. While proteins absorb at 280 nanometers (nm), DNA absorbs at 260 nm. Thus, the samples should generate a large peak at 260 nm. 1 OD ~ 50 μg/mL double-stranded DNA.[8] The Nanodrop will provide a 260/280 ratio to measure potential protein contamination. At the minimum, the A260/A280 ratio for DNA should be ~1.8, since values less than 1.7 tend to generate less than optimal results [26].

1) Rinse Nanodrop with a few drop of nuclease-free water to clean the reader.

2) Rinse with 2 μL of nuclease-free water when prompted to load water sample for a standard. It is not necessary to blank the instrument between the standard and unknown samples measurements.
3) Run 2 μL of each sample. After each sample rinse with 2 μL of nuclease-free water. Clean off top and bottom arm of reader with Kimwipes after each sample.

ExoSAP-IT® PCR Product Cleanup

Affymetrix ExoSAP-IT® PCR Product Cleanup treats PCR products by removing unused primers and nucleotides. This step prepares the PCR products for use in DNA sequencing by eliminating PCR primer carryover and other amplification products [27].

1) Remove ExoSAP-IT® reagent from -20°C freezer, thaw, and keep on ice throughout this procedure. Thoroughly mix reagent by light vortexing for a few seconds as various components could have settled out or been distributed unevenly, which will affect the precision of the samples. Lightly centrifuge for a few seconds to ensure all contents are at the bottom before pipetting.
2) Mix 5 μL of a post-PCR reaction product with 2 μL of ExoSAP-IT® reagent for a combined 7 μL reaction volume. When treating PCR product volumes greater than 5 μL, simply increase the amount of ExoSAP-IT® reagent proportionally.
3) Incubate at 37°C for 15 minutes to degrade remaining primers and nucleotides in PCR machine.
4) Incubate at 80°C for 15 minutes to inactivate ExoSAP-IT® reagent.
5) The PCR product is now ready for use in DNA sequencing. The treated PCR products may be stored at -20°C until required.

HIGH-THROUGHPUT SEQUENCING

DNA sequencing is used to verify that the PCR products amplified through PCR and shown in the gel electrophoresis are the desired sequences. Cornell University's Institute of Biotechnology contains the Biotechnology Resource Center, which offers Sanger sequencing as a service. DNA sequencing is performed using Applied Biosystems Automated 3730xl DNA Analyzers, Big Dye Terminator chemistry and AmpliTaq-FS DNA Polymerase [26]. The steps for preparing the PCR products whose DNA concentration have been found adequate and that have been cleaned up with ExoSAP-IT for high-throughput sequencing are as follows [26]:

1) For Sanger Sequencing primer submission, orders must be placed using the online ordering system. Select "DNA Sequencing" for "Facility" and "Single Tube Sequence" as the "service".
2) Place the required amount of PCR product and 8 pmole of primer in the specified vials and bring up to 18 μL. To determine the required amount of PCR product to add, use the following formula:

   #base pairs/5.0 = amount of PCR product in ng needed.

   nosZ long amp: 706 bp/5.0 = 141.2 ng of DNA

   norB long amp: 542 bp/5.0 = 108.4 ng of DNA
3) Samples are accepted in 500 μL standalone screw top vials, in USA Scientific (catalog number 1405-9700 to 1405-9706 select color, 1405-9799-mixed color caps).
4) The turnaround time for receiving the sequencing results is around 1-2 days. The sequencing reaction to be analyzed will undergo 5 stages: cycle sequencing, post-sequencing cleanup, capillary electrophoresis, data collection , and sequence analysis.

CREATING DEGENERATE PRIMERS STANDARDS

Quantitative PCR standards of the nosZ and norB degenerate primers must be generated so that when testing soil samples, their cycle threshold (C_T) can be compared to a predetermined scale of DNA gene copies in a series of 10^8-10^1 dilutions in order to determine their approximate DNA concentration. The cycle threshold is the number of cycles required for the fluorescent signal of the SYBR Green Supermix to cross a predetermined background level, or threshold.

QUANTITATIVE PCR

1) Turn on Bio-Rad iCycler Thermal Cycler with iQ5 Multicolor Real-Time PCR Detection System and optical lamp at least 30 minutes in advance of running a qPCR plate.
2) To avoid DNA cross-contamination, clean the set-up area and equipment to be used with 10% bleach. Thaw Bio-Rad iQ SYBR Green Supermix, nosZ and norB DNA sample templates, forward and reverse primers, and standards on ice, keeping SYBR Green Mix out of the light. Thoroughly mix all reagents during the set-up procedure by light vortexing for a few seconds as various components could have settled out or been distributed unevenly, which will affect the precision of the samples. Lightly centrifuge for a few seconds to ensure all contents are at the bottom before pipetting.
3) In order to create the 10^8-10^1 Dilutions of the samples, in this case the *Paracoccus denitrificans* (PD 1222) PCR samples run using the degenerate primers, the original copy number of the samples must be determined as seen in Table 9. This calculation assumes that the mass of a base pair in double stranded DNA is 660 Da, or 660 g in weight [28]. The formula weight of the double strand of nucleotides can be determined by multiplying the length of the nucleotides in base pairs by the average formula weight of the types of

nucleotides in the strand. The inverse of the formula weight is the number of moles of DNA template found in one gram. The number of molecules of the template per gram can thus be calculated using Avogadro's number, 6.022×10^{23} molecules/mole. Then, the number of molecules, also known as the number of copies of template in the sample, can be estimated by multiplying by 10^9 to covert to ng and then multiplying by the amount of template in ng.

$$Number\ of\ copies = \frac{sample\ amount\ (ng) \times \frac{6.022 \times 10^{23}\ molecules}{mole}}{DNA\ template\ length\ (bp) \times \frac{1 \times 10^9\ ng}{g} \times \frac{660\ g}{mole\ of\ bp}}$$

STANDARD	COPY NUMBER	STOCK NEEDED (µL)	NUCLEASE-FREE WATER (µL)
Original	8.34×10^{10}		
10^8	8.34×10^8	1.5	148.5
$10^{7\text{-}1}$	$8.34 \times 10^{7\text{-}1}$	15	135

Table 8: NorB 10^1-10^8 DNA Template Dilutions.

STANDARD	COPY NUMBER	STOCK NEEDED (µL)	NUCLEASE-FREE WATER (µL)
Original	8.63×10^{10}		
10^8	8.63×10^8	1.5	148.5
$10^{7\text{-}1}$	$8.63 \times 10^{7\text{-}1}$	15	135

Table 9: NosZ 10^1-10^8 DNA Template Dilutions.

4) Create a "Master Mix" for each primer set as seen in Table 10, combining SYBR Green reagent, forward and reverse primers, and ultra-pure water according to the Vortex thoroughly. Make 1-2 samples more worth of Master Mix. One tube of SYBR Green contains enough for 37 samples total.

REAGENT STOCK	AMOUNT (µL) PER 25 µL REACTION
Bio-Rad iQ SYBR Green Supermix	12.5
IDT Forward Primer (15 µM)	1.167
IDT Reverse Primer (15 µM)	1.167

DNA Template (50 ng) <u>nosZ and norB DNA Template Samples Tested</u> • Pure Culture PD1222 (*Paracoccus denitrificans)*: 16.6 ng/µL	Varies depending on pure culture concentration: • PD1222: $50\text{ ng} \times \frac{\mu L}{16.6\text{ ng}} = 3\ \mu L$
Promega Corporation Nuclease-Free Water	Varies depending on DNA template amount: • PD: 7.17

Table 10: NosZ and norB Master mix components and amounts for qPCR.

5) For each triplicate sample, distribute 73 µL Master Mix into 0.2 mL PCR tubes. Do not let the drip on the end of the pipette tip drip into the tube.
6) Add 10 µL cDNA to each tube.
7) Making sure that the sample is adequately mixed, vortex 1 second. Pipette up and down gently 6 times, distributing 25 µL of each sample into a well in the Bio-Rad iQ 96-well PCR plate. Be careful not to expel tip all the way so as not to create bubbles.
8) To create no template control (NTC) samples to ensure no background amplification from environmental contaminants or primer dimers occurs, substitute the DNA templates being tested with nuclease-free water for each dilution.
9) Seal plate with Bio-Rad Microseal "B" Adhesive Seals. .
10) Centrifuge plate for 15 seconds at maximum speed, around 10,000 rpm, in order to remove any bubbles or drops on the sides of the wells. Recheck for bubbles. Flick the bottoms of the tubes to remove bubbles, but remember to re-centrifuge the plate after doing this.
11) Run plate on iCycler PCR machine using protocols as seen in Table 11 and Table 12.

Cycle	TEMPERATURE (°C)	TIME (min:sec)	CYCLES
1	55	01:00	1
	95	05:00	1
2	95	00:45	40
	55	00:45	40
	72	00:45	40
	Data collection enabled.		
3	60	00:10	70
	Increase setpoint temperature after cycle 2 by 0.5°C. Melt curve data collection and analysis enabled.		
4	4	HOLD	HOLD

Table 11. Tentative nosZ degenerate primers qPCR thermal cycling program.

Cycle	TEMPERATURE (°C)	TIME (min:sec)	CYCLES
1	55	02:00	1
	95	05:00	1
2	95	00:30	40
	55	00:30	40
	72	01:00	40
	60	00:30	40
	Data collection enabled.		
3	60	00:10	80
	Increase setpoint temperature after cycle 2 by 0.5°C. Melt curve data collection and analysis enabled.		
4	4	HOLD	HOLD

Table 12. Tentative norB degenerate primers qPCR thermal cycling program.

12) Come back after 2.5 hours to turn of iCycler and optical lamp or bulb may burn out more quickly.

DISCUSSION

VALIDATING DEGENERATE PRIMERS

The set of degenerate F and R primers for norB and nosZ were designed and validated successfully. Gel electrophoresis indicated a strong band at the desired amplicon sizes, the Nanodrop showed a sizably DNA concentration with limited contamination, and high-throughput sequencing returned clean electropherograms. Thus, the degenerate primer sequences as seen in Table 12 will be used for future studies.

VALIDATING LONG AMPLICON PRIMERS

END-POINT PCR

After testing for the ideal range of annealing temperatures and analyzing the results on gel electrophoresis and high-throughput sequencing, the optimal annealing temperature for the long amplicon norB and nosZ primers was determined. As stated before, it is important to identify the ideal annealing temperature in order to maximize the amount of PCR product generated from the limited amount of original soil samples obtained, resulting in the adjusted PCR protocols as seen below in Table 12 and 13.

CYCLE STEP	TEMPERATURE (°C)	TIME (min:sec)	CYCLES
Initial Denaturation	95	10:00	1
Denaturation	95	00:30	30
Annealing	54	00:45	30
Extension	72	00:45	30
Final Extension	72	05:00	1
HOLD	4	HOLD	HOLD

Table 12. Optimized long amplicon norB primer PCR thermal cycling program.

CYCLE STEP	TEMPERATURE (°C)	TIME (min:sec)	CYCLES
Initial Denaturation	95	10:00	1
Denaturation	95	00:30	30
Annealing	55	00:45	30
Extension	72	00:45	30
Final Extension	72	05:00	1
HOLD	4	HOLD	HOLD

Table 13. Optimized long amplicon nosZ PCR thermal cycling program.

GEL ELECTROPHORESIS

After testing the 1st set of F and R nosZ long amplicon primers for their validity, it was ascertained that they were unable to be used. Although the primers were designed to encompass the nosZ primer target sequence, gel electrophoresis showed that the amplicon failed to be a strong band. In addition, there was evidence of additional amplicons of smaller sizes being generated, which indicated there was non-specific binding occurring. While results from the Nanodrop showed an adequate DNA concentration, high-throughput sequencing returned noisy data. As a result, a 2nd set of primers nosZ long amplicon F and R needed to be designed and validated. After gel electrophoresis indicated a strong band at the desired amplicon size, the Nanodrop showed a sizably DNA concentration with limited contamination, and high-throughput sequencing returned a clean electropherogram, it was determined that the 2nd set nosZ long amplicon F and R will be used for further studies as seen in Table 14.

PRIMER	5' TO 3' SEQUENCE	AMPLICON SIZE (base pairs)
nosZ long amp F 2nd set	CCTGTTCACCCGCTATATCC	706
nosZ long amp R 2nd set	GTAGACCCGCACCTTGTT	706
norB long amp F	TGCTGATGGGCCTTTGG	542
norB long amp R	GCCATAGAAGGCCAGGTG	542

Table 14. Long Amplicon denitrification pathway primers.

Because the gel electrophoresis protocol called for the use of GelRed™, there may be some instances when results are uncertain, as was the case when the nosZ long amp F and R 1st

set failed to show clear band results. To be aware of whether it is the fault of the GelRed™ or of the primers themselves, it is important to rule out the possible situations GelRed™ could or could not be responsible for when troubleshooting. In cases where there are smeared DNA in a precast gel, it may be a result of overloading, which is typically the case with DNA ladders since GelRed™ is more sensitive than the typical protocol which uses EtBr. If even after reducing the amount of DNA loaded by one-half to one-third doesn't work, the gel may need to be post-stained in a solution of GelRed™ 10,000Z stock solution diluted 3,300 fold to make a 3X staining solution in about 50 mL of water for a small minigel. The bands should form brightly after about thirty minutes of agitation in room temperature. Larger DNA fragments may require pouring a lower percentage of agarose gel for better resolution. In cases where there are discrepant DNA migration in the pre-cast gel, because GelRed™ is designed to be a larger molecule to prevent it from entering cells, the migration of DNA may be affected by the dye: DNA ratio. To combat this phenomenon, it may be necessary to reduce the amount of DNA loaded by one-half to one-third, reduce the amount of dye used, or post-stain the gel to avoid interference the dye may have had on migration during electrophoresis. Lastly, if the bands display weak fluorescence, decreased dye performance over time, or a film of dye remains after post-staining, it is possible the dye has precipitated out of the solution. Thus, GelRed™ should be heated to 45-50° C for two minutes and vortexed to redissolve, keeping it stored at room temperature [23].

If GelRed™ was determined to not be the reason for undesirable bands, then are a number of problems that could be troubleshooted from gel electrophoresis. Low intensity of the DNA bands, as was the case with the 1st set of F and R long amplicon nosZ primers, could have been due to insufficient or uneven staining, insufficient amount of DNA bands generated, DNA

diffusion from long term storage of the gel, the DNA running off the gel, or the DNA being masked by the loading dyes. In cases where there are smeared DNA bands, it could be a result of DNA degradation by nucleases, improper electrophoresis conditions, excess DNA loaded, or possibly poorly formed gel wells. Atypical banding patterns could be a result of denatured DNA, different loading conditions the samples and the ladder DNA, improper electrophoresis conditions, incorrect gel percentage or running buffer, or even high salt concentrations in the sample. In the case of curved DNA bands, the gel is incompletely immersed in electrophoresis buffer, low sample volume, improper electrophoresis conditions, or bubbles and physical particles in the gel wells or in the gel. DNA remains in the gel could be a result of poorly formed gel wells, excess DNA loaded, contamination of the DNA sample, or even the gel shift effect, which occurs if the presence of DNA binding proteins in the sample, such as ligases, phosphatases, or restriction enzymes, alter DNA migration in the gel and cause the DNA to remain in the gel wells. Finally, incorrect quantification data could be a result of different loading conditions for the sample and ladder DNA, incorrect ladder band chosen for the quantification of the sample, improper quantification methods used, uneven staining of the gel, or DNA masking by electrophoresis tracking dyes [29].

HIGH-THROUGHPUT SEQUENCING

Successfully interpreting the electropherograms, or plot of results of the DNA sequences, obtained from high-throughput sequencing is vital for determining what is good, reliable sequence data, as one needed to do in order to rule out using the 1st set of nosZ long amplicon F and R primers. The electropherograms should indicate that the strongest signals (highest peak) of each base pair correlates with the predetermined PCR product sequence, with the peaks evenly

spaced [26]. Ideally, peak height towards the middle of the sequence will also be significantly higher than earlier fragments, with little or no background interference at the peak baselines. Unreliable data is any peak which is jumbled, rounded, short, or not well resolved, which the sequencing analysis programming assigning an "N" to a base when it can't identify a single nucleotide in a position. In addition, the primer sequences of the data will not be included in the trace results, meaning that theoretically, the first nucleotide of the sequence should be identical to the one after the first nucleotide of the primer although this is not often the case since the least dependable areas of sequence are at the beginning and end.

The most common causes of failure of good sequence data tends to relate back to the concentration and purity of the original template DNA. There may be other cases when there is no sequence data, in which the peaks are short and muddled, which may be a result of a lack of reliable priming site, not enough DNA/primer in the reaction tube, inhibitory contaminants, or even possibly expired reagents. When data has noisy, weak signals, it may be due to not enough DNA, inhibitory contaminants such as salts or phenol, degraded DNA from nucleases, excessive UV light exposure, bisulfite treatment, repeated freeze-thaw, or simply such inefficient primer binding. In cases where there are multiple peaks within the sequence at the beginning, this could be a result of multiple priming sites in PCR, where they can hybridize to more than one position on the template DNA, giving rise to multiple PCR products. The visualization of the PCR products on the agarose gel should have indicated this, which may necessitate gel purification of the desired product, although PCR redesign is preferred. Other factors for this may include the primers acting as both forward and reverse one, residual PCR primers and/or dNTPs primers with n-1 population, or primers with high Tm, which often suggest that there is high GC content. In cases where there are multiple peaks within the sequence at the middle, this could indicate

frame shift mutations, homopolymeric regions, or compression. Truncated sequences could indicate secondary structure and/or abnormal sequence motifs, linearized DNA, too much DNA, salts, or repetitive regions. Lastly, sequencing artifacts such as dye blobs, loss of resolution or spikes which may indicate an instrument or post-cycle sequencing cleanup issue [30].

CREATING DEGENERATE PRIMERS STANDARDS

QUANTITATIVE PCR

Although degenerate primer standards for nosZ and norB were generated, the standard curves shown in Figure 11 and Figure 12 are not the final standard, because their qPCR protocols still need to be adjusted for maximum efficiency of the templates. Because PCR amplification efficiency is that rate at which a PCR amplicon is generated, expressed as a percentage value, an accurate estimate of PCR efficiency depends on a variety of reagent, experimental set-up, sample quality, and analysis factors [31]. This implies that if a PCR amplicon has doubled in quantity, then then PCR assay has 100% efficiency. A standard curve is represented as a semi-log regression line plot of C_T values vs. log of DNA concentration of the original DNA template. Because the samples were tested in triplicate, the points should all be approximately the same value, generating a R^2 value of at least 0.99. Any values determined to be less than 0.99 could be a result of poor pipetting of standard identical replicates or standards, and potentially degraded DNA template [32]. However, even with a potentially good R^2 value, the slope of the standard curve may still be inaccurate, due to consistent pipetting excess of the serial dilutions which may lead to a perceived lower PCR efficiency of the assay, or a consistent pipetting deficit of the serial dilutions/standard sample which may lead to a perceived higher PCR efficiency of the

assay. Similarly, consistent pipetting excess of the standard sample may lead to a perceived higher PCR efficiency of the assay, or a consistent pipetting deficit of the standard sample which may lead to a perceived lower PCR efficiency of the assay [31].

Therefore, the ideal efficiency of amplification of this control template should be between 90-110% with a slope of -3.58 to -3.10. A PCR reaction with 100% efficiency would have a slope of -3.32 [32]. In other words, a 100% efficient reaction yields a 10-fold increase in PCR amplicon every 3.32 cycles during the exponential phase of amplification, where $\log_2 10 = 3.32$. Slopes more negative than -3.32 indicate reactions are less than 100% efficient, while slopes more positive than -3.32 may indicate sample quality or pipetting problems. If one wants to estimate the efficiency of their assay, they can use the calculation Efficiency = $(10^{-1/\text{slope}} - 1) \times 100$ to assess whether standard needs to be redone [32]. In addition, the NTC wells should be checked for background amplification from environmental contaminants or primer dimers. Results are deemed acceptable if the C_T value is greater than 10 cycles after the last control template amplification. Afterwards, the threshold fluorescence should be locked in on the standard curve [31].

From just the data on the number of C_T each sample created, the qualitative amount of target nucleic acid can be determined. The C_T levels are inversely proportional to the amount of target nucleic acid in the sample, meaning that the lower the C_T count, the greater the amount of DNA in the sample. Thus, $C_T < 29$ are strong positive reactions indicative of abundant target DNA, while C_Ts of 30-37 are positive reactions indicative of moderate amounts of target nucleic acid, and C_Ts of 38-40 are weak reactions indicative of minimal amounts of target nucleic acid which could represent an infection state or environmental contamination. To attain a true quantitative measurement of the amount of DNA present in each sample, the best-fit linear line

of the standard points should be plotted and the y = mx + b equation for the line determined. By plugging y = C_T value into the equation, the x = DNA concentration can be estimated [31].

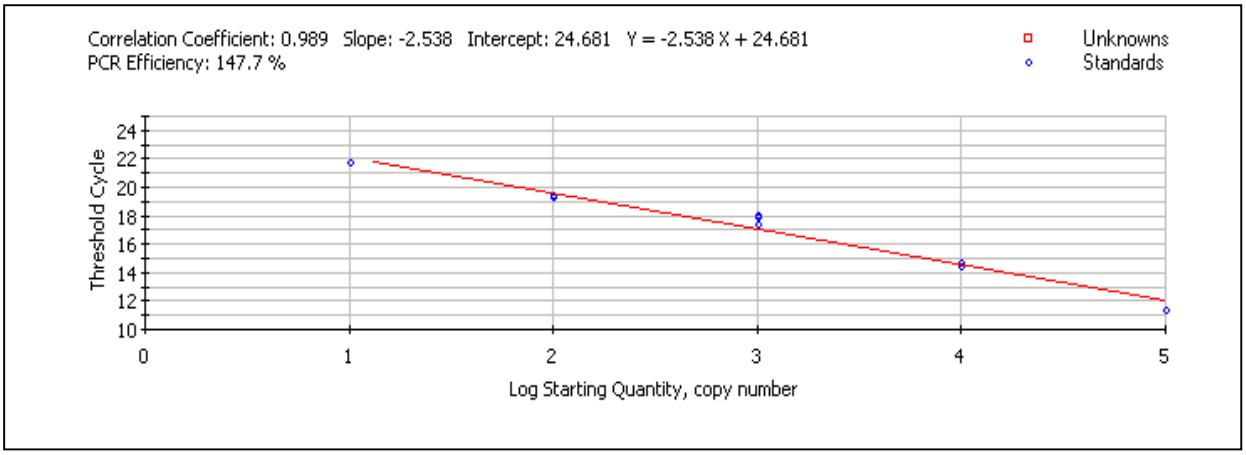

Figure 11. Tentative nosZ degenerate primers standard curve.

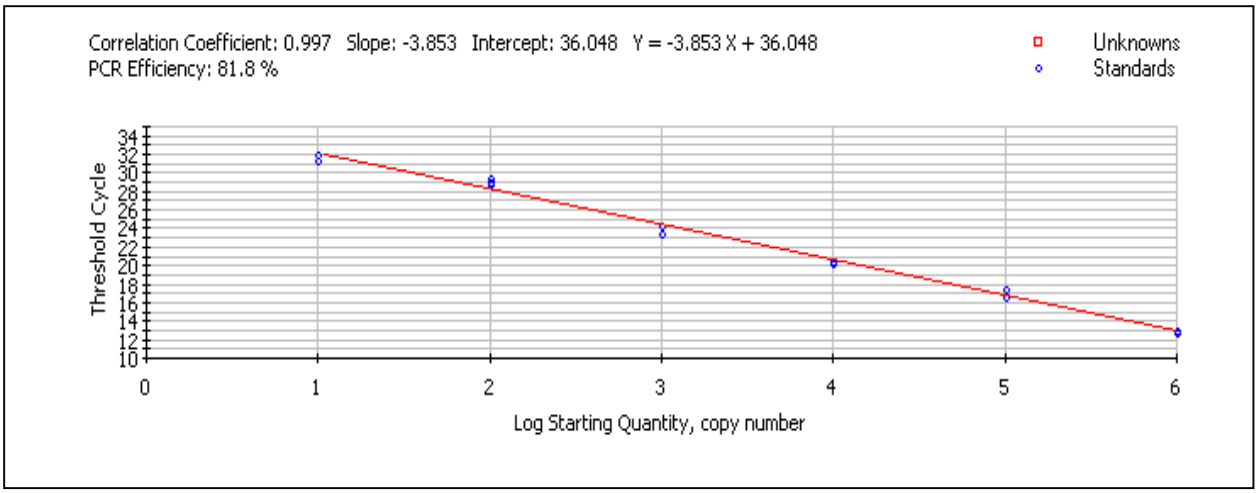

Figure 12. Tentative norB degenerate primers standard curve.

FUTURE WORK

The design and optimization of the qPCR and PCR forward and reverse norB and nosZ primers and their protocols will be used to verify the initial hypothesis that soil depth is a greater predictor of denitrifier activity than the presence of a septic leachfield. As the soils samples across the nine sites continue to undergo qPCR tests for the presence of nosZ and norB producing organisms, their results will be used to corroborate some of the preliminary results of the atmospheric gas flux measurements.

The second objective, to determine the microbial community composition of the nine sites, will be addressed by employing high-throughput sequencing techniques. The goal is to gain a comprehensive understanding of the microbes involved in GHG cycling from septic leach fields. As suggested in previous studies, further studies may be conducted to capture the diversity of nosZ-using organisms [13]. This could be done by surveying microbial communities at additional sites and using other primers designed to target atypical nosZ. In addition, because the samples that will be analyzed for this overall study were taken during one season in one year, it may be worthwhile to collect samples over a longer period of time to showcase any seasonal differences in the microbial community composition. There is the potential for differences between dry and submerged soils to occur, as microbial activity increases during warmer months of the year [18].

CONCLUSION

The eventual results of this overall study will help shed lights on the roles denitrifiers, methanogens, and methanotrophs play in the cycling of the potent greenhouse gases N_2O and CH_4 to net fluxes from septic system leachfields. The atmospheric flux data from the earlier half of the study will help corroborate the molecular biomarker work using the primer sets targeting *mcrA*, *pmoA*, *nosZ*, and *norB* to quantify the microbes' presence and abundance. Eventually, high throughput sequencing of the soil samples will identify the microbial communities primarily involved in greenhouse gas cycling. The results of this study will allow for better investigation into the mitigation of potential greenhouse gases from septic systems, and allow for broader knowledge of the rapidly emerging technology of molecular biology in use for analyzing environmental systems. In the future, the knowledge from this and similar studies will aid in the design of measures implemented to resolve their potentially negative impacts.

REFERENCES

[1] Westco Group Incorporated. A Leader in Septic System Drain Field Management and Restoration. Typical Septic System Overview. Accessed June 2016. http://www.westcogroupinc.com/wp-content/uploads/2016/06/septic_layout.png

[2] U.S. Environmental Protection Agency. Office of Wastewater Management. Case Studies of Individual and Clustered (Decentralized) Wastewater Management Programs: States and Community Management Programs. (June 2012) https://www.epa.gov/sites/production/files/2015-06/documents/decentralized-case-studies-2012.pdf

[3] Understanding Your Septic System. Houselogic.com. (2016) https://www.houselogic.com/organize-maintain/home-maintenance-tips/understanding-your-septic-system/.

[4] How Your Septic System Works. https://www.epa.gov/septic/how-your-septic-system-works.

[5] Diaz-Valbuena, L. R., Leverenz, H. L., Cappa, C. D., Tchobanoglous, G., Horwath, W. R., & Darby, J. L. (2011). Methane, carbon dioxide, and nitrous oxide emissions from septic tank systems. *Environmental Science & Technology*, *45*(7), 2741–7.

[6] Luton, P.E., Wayne J.M., Sharp R.J., Riley P.W. (2002). The mcrA gene as an alternative to 16S rRNA in the phylogenetic analysis of methanogen populations in landfill. *Microbiology*.148(Pt 11):3521-30.

[7] Steinberg, L. M., & Regan, J. M. (2009). mcrA-targeted real-time quantitative PCR method to examine methanogen communities. *Applied and Environmental Microbiology*, *75*(13), 4435–42.

[8] Bourne, D. G., McDonald, I. R., & Murrell, J. C. (2001). Comparison of pmoA PCR primer sets as tools for investigating methanotroph diversity in three Danish soils. *Applied and Environmental Microbiology*, *67*(9), 3802–9.

[9] Costello, A. M., & Lidstrom, M. E. (1999). Molecular characterization of functional and phylogenetic genes from natural populations of methanotrophs in lake sediments. *Applied and Environmental Microbiology*, *65*(11), 5066–74.

[10] Kolb, S., Knief, C., Dunfield, P. F., & Conrad, R. (2005). Abundance and activity of uncultured methanotrophic bacteria involved in the consumption of atmospheric methane in two forest soils. *Environmental Microbiology*, *7*(8), 1150–61.

[11] Henry, S., Bru, D. Stres, B. Hallet S., Philippot, L. (2006) Quantitative Detection of the nosZ Gene, Encoding Nitrous Oxide Reductase, and Comparison of the Abundances of 16S rRNA, narG, nirK, and nosZ Genes in Soils. Appl Environ Microbiol. 72(8): 5181-5189.

[12] Richardson, D., Felgate, H., Watmough, N., Thomson, A., & Baggs, E. (2009). Mitigating release of the potent greenhouse gas N(2)O from the nitrogen cycle - could enzymic regulation hold the key? *Trends in Biotechnology*, *27*(7), 388–97.

[13] Sanford, R.A., D.D. Wagner, Q. Wu, et al. (2012). Unexpected nondenitrifier nitrous oxide reductase gene diversity and abundance in soils. *Proc. Nat. Acad. Sci.*, 109 (48): 19709-19714.

[14] Philippot, L. (2002). Denitrifying genes in bacterial and Archaeal genomes. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression*, *1577*(3), 355–376.

[15] Shapleigh, J. P. (2006). Denitrifying Prokaryotes: Defining the Denitrifiers. *The Prokaryotes*. 769–792.

[16] Alvarez L., Bricio C., Blesa A., Hidalgo A., Berenguer J. (2014). Transferable Denitrification Capability of Thermus thermophiles. *Appl. Environ. Microbiol.* 80(1), 19-28.

[17] Fernandez-Baca, C. CSBC Small Grant Fall 2014. *Small Grant Septics Research Proposal*. Cornell University.

[18] Hegarty, B. (2013) *Quantitative Detection of Denitrification Genes—With A Particular Focus on Atypical Nitrous Oxide Reductase Diversity*. Senior Honors Thesis. Cornell University.

[19] Braker, G., Tiedje, J.M. (2003) Nitric oxide reductase (norB) genes from pure cultures and environmental samples. Applied Environ. Microbiol. 69, 3476-3483.

[20] Ruth E. Richardson Lab Protocol. Endpoint PCR with GoTaq DNA Polymerase. Cornell University.

[21] Ruth E. Richardson Lab Protocol. Analysis of DNA by Agarose Gel Electrophoresis. Cornell University.

[22] 0.5X TBE Electrophoresis Buffer. Accessed July 2016. https://www.eeb.ucla.edu/Faculty/Barber/Web%20Protocols/Protocol7.pdf

[23] Biotium: Glowing Products for Science™. GelRed™ Nucleic Acid Gel Station, 10,000X in Water Product Information. http://biotium.com/wp-content/uploads/2013/07/PI-41003.pdf

[24] Integrated Microbial Genomes & Microbiomes (IMG/M). https://img.jgi.doe.gov/cgi-bin/m/main.cgi

[25] Integrated DNA Technologies (IDT). https://www.idtdna.com/pages/products/dna-rna/custom-dna-oligos.

[26] Cornell University. DNA Sequencing Handbook. Revised October 2011. http://www.biotech.cornell.edu/sites/default/files/uploads/SeqHandbook.pdf

[27] Affymetrix USB. USB® ExoSAP-IT® PCR Product Cleanup Brief Protocol. Revised December 15. http://media.affymetrix.com/support/technical/usb/brief_proto/78200B.pdf

[28] SciencePrimer.com Copy Number Calculator for Realtime PCR. Andrew Staroscik 2011-2016. http://scienceprimer.com/copy-number-calculator-for-realtime-pcr

[29] Troubleshooting Guide for DNA Electrophoresis. ISO 9001. 9. DNA Electrophoresis: Protocols and Recommendations for DNA Electrophoresis. www.fermentas.com.

[30] Interpreting Electropherograms. University of Delaware. http://www1.udel.edu/dnasequence/Site/Interpreting_Electropherograms.html

[31] Guide to Performing Relative Quantitation of Gene Expression Using Real-Time Quantitative PCR. Applied Biosystems. 2004. http://www6.appliedbiosystems.com/support/tutorials/pdf/performing_rq_gene_exp_rtpcr.pdf

[32] qPCR Quantification Protocol Guide. Illumina. September 2009. http://www.bu.edu/iscf/files/2010/07/qPCR_Quant_Protocol_Guide_11322363_A.pdf